DNA-Based Computing & Development of Efficient Operators

for Solving Traveling Salesman Problem

Tai Hyun Park

School of Chemical Engineering

Seoul National University

Contents

- Introduction

- DNA computing for solving traveling salesman problem
  - Traveling salesman problem (TSP)
  - Molecular algorithm for TSP
  - Experimental implementation

- Unit operators in DNA computing
  - Denaturation temperature gradient polymerase chain reaction (DTG-PCR)
  - Initial pool generation using parallel overlap assembly (POA)

Rise and Growth of DNA Computing

- Adleman’s work in 1994
  - Hamiltonian path problem (graph problem, NP problem)
  - City and road information representation using DNA sequences (indicative information)
  - Solution-based DNA computing
- Liu’s work in 2000
  - SAT problem
  - Surface-based DNA computing
- Benenson’s work in 2004
  - Application to disease diagnosis and drug (antisense) administration

Our Work in DNA Computing Era

- Solving traveling salesman problem
  - Graph problem with weighted edges
  - Representation of numerical information
  - Expansion to larger problems (in progress)
- Development of unit operators
  - Denaturation temperature gradient PCR (DTG-PCR)
  - Initial pool generation for TSP using parallel overlap assembly (POA)
  - Modeling and simulation of DTG-PCR and POA
- New application to medical diagnosis
  - Disease diagnosis model development
  - Experimental verification (in progress)

Objective (traveling salesman problem)

- Application of DNA computing to mathematical problems
  - 7-city traveling salesman problem
  - Graph problem with weighted edges

- Representation of numerical values (weighted edges)

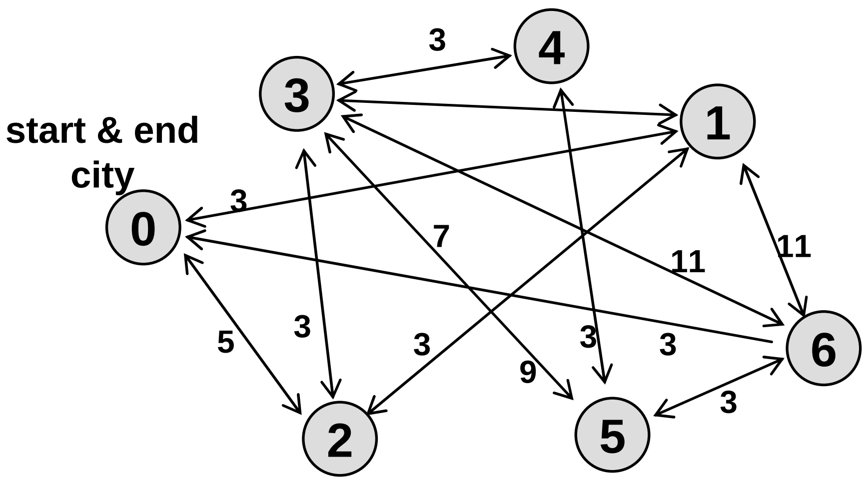
- Expansion to practical problems
  - 26-variable TSP ← 26 subway stations in Seoul

Traveling Salesman Problem

- Find...
  - The cheapest way of visiting all the cities and returning to the starting point

    when a number of cities to visit and the traveling cost between each pair of cities are given.
- Previous work for weight (cost) representation
  - DNA length
  - DNA concentration
- Our method for weight (cost) representation
  - Thermal stability of DNA duplex
  - Melting temperature (T_m), GC content

Target Problem & Encoding Method

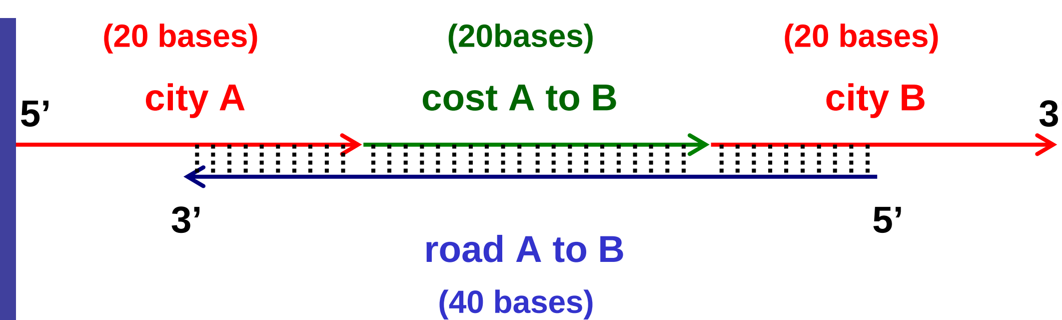


7-city traveling salesman problem

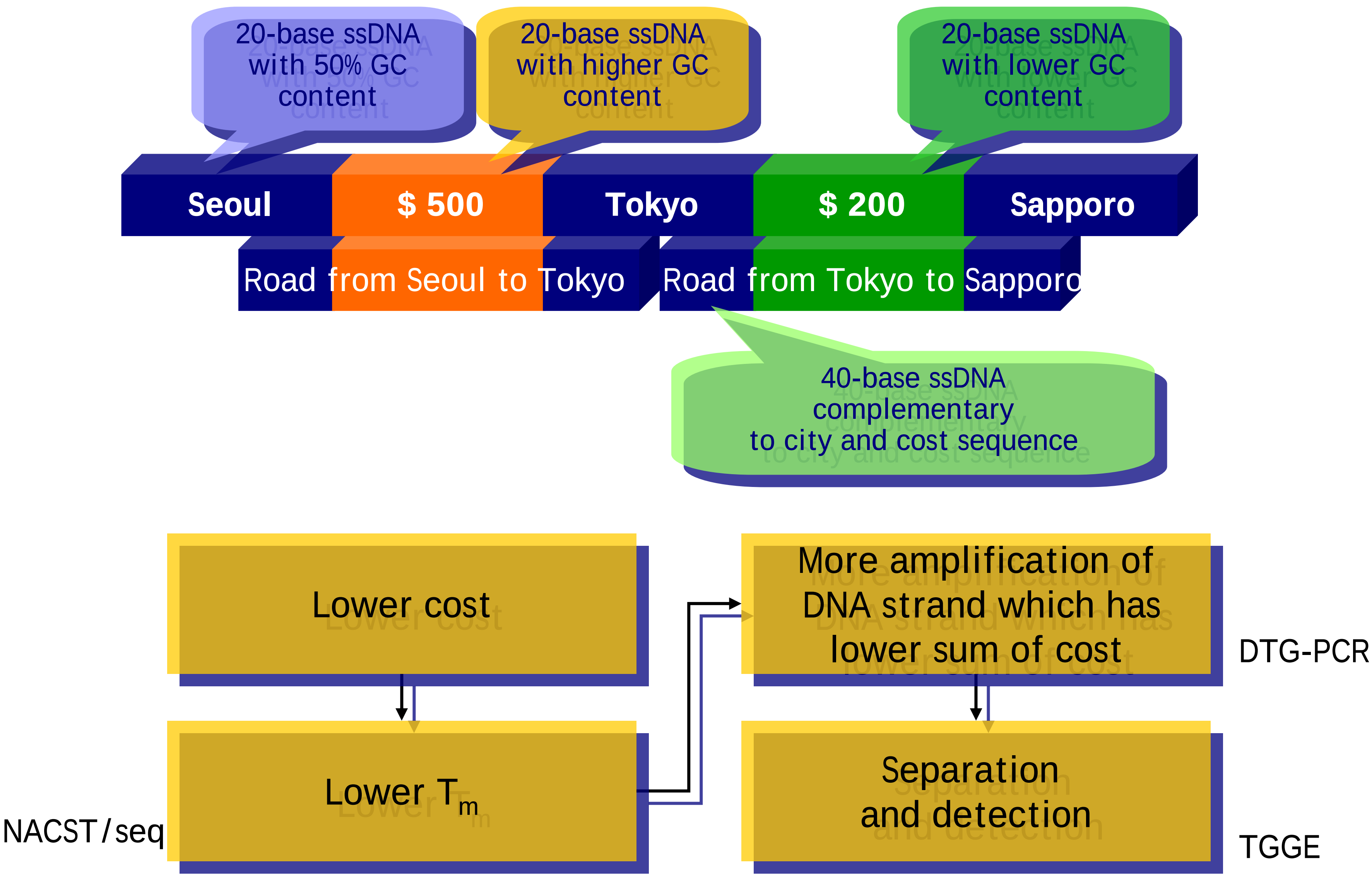
- 7 cities (0 to 6), 23 roads, 5 costs
- optimal path: ‘0→1→2→3→4→5→6→0’

Oligonucleotides & Encoding

- cities and costs are 20-base ssDNA
- roads are 40-base ssDNA
- 35 oligonucleotides (7 cities, 23 roads, 5 costs)



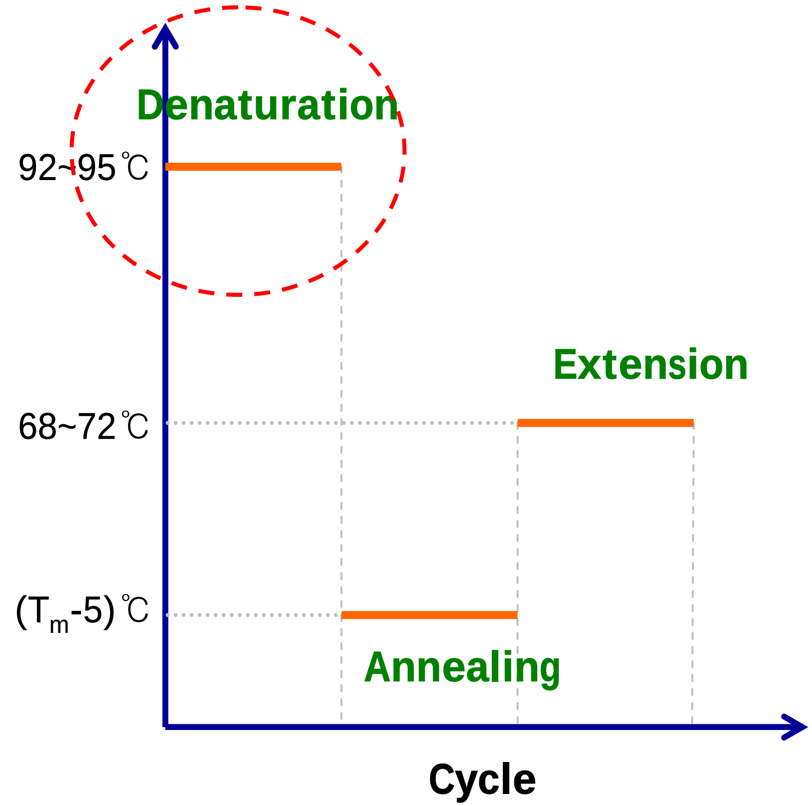
Weight (Cost) Encoding



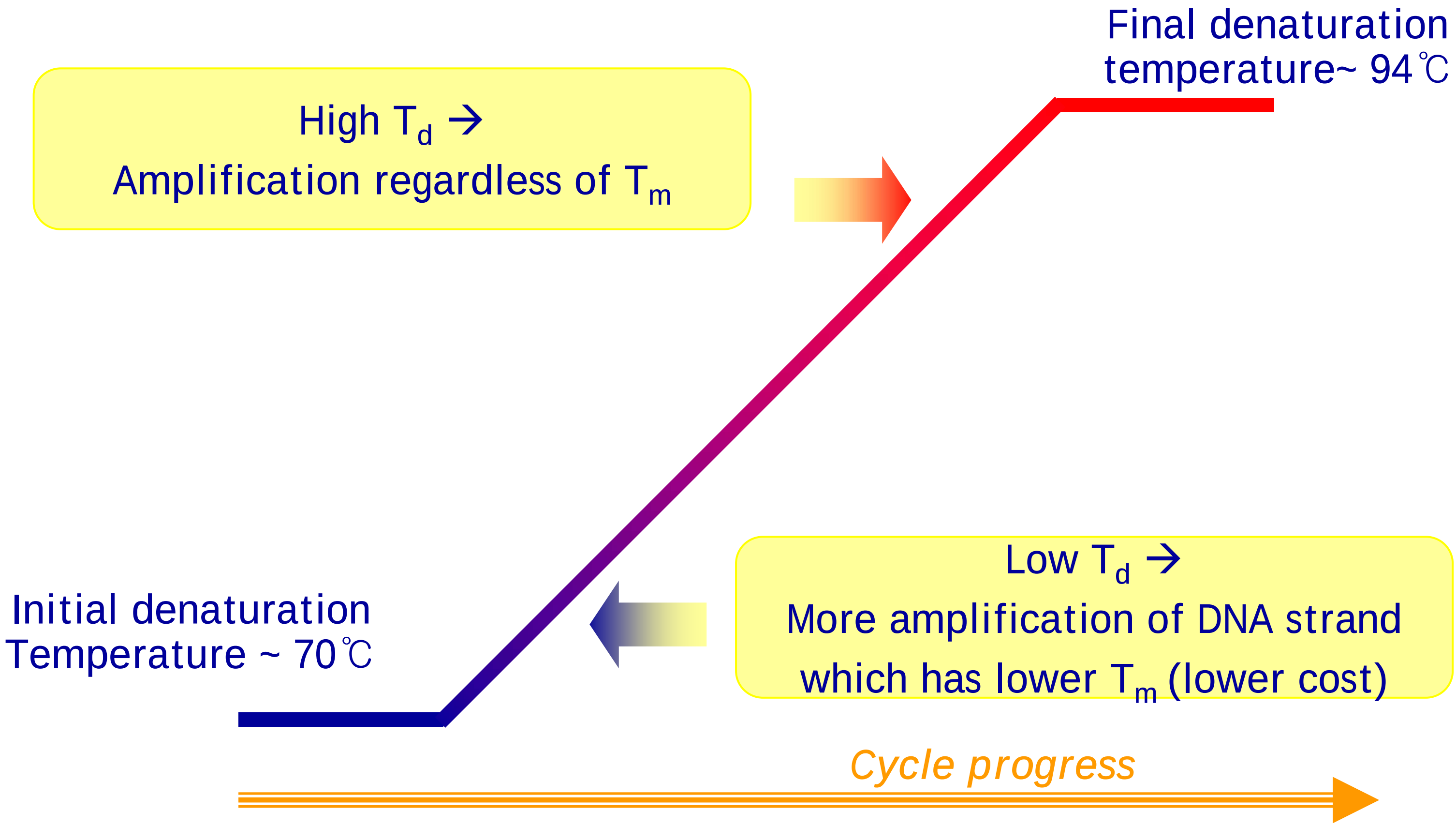
Denaturation Temperature Gradient Polymerase Chain Reaction (DTG-PCR)

- Conventional PCR
  - Denaturation (T_d)
  - Annealing (T_a)
  - Extention (T_e)

- Modification of conventional PCR protocol
  - Variation in T_d

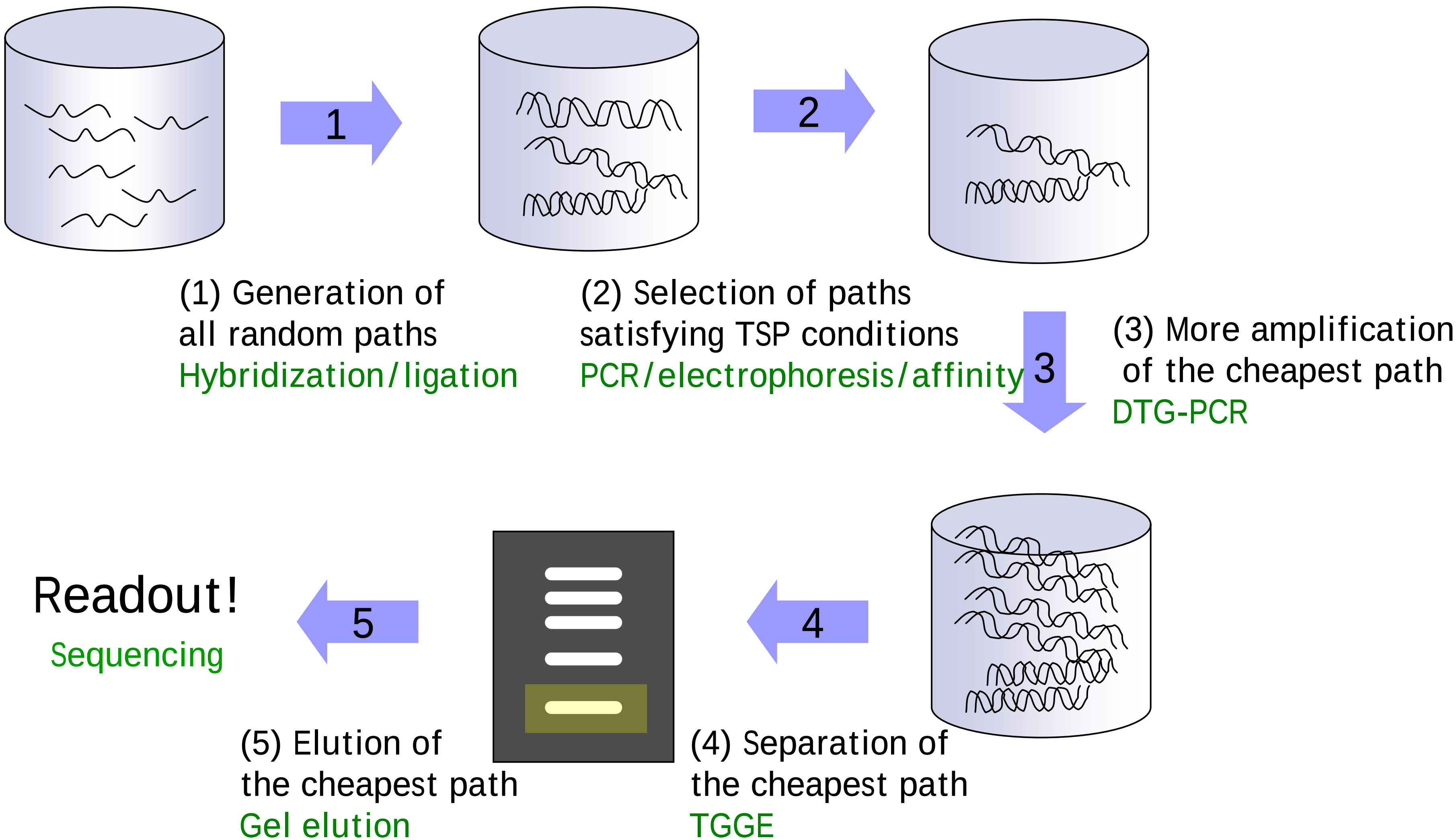


Denaturation Temperature Gradient



- Biased operator: more amplification of DNA strands with lower T_m
  → biased search for lower cost

Molecular Algorithm



Sequence Design for Cities and Costs

- Using NACST/seq
- Non-cross hybridization
- Similar T_m among cities
- Different T_m among costs

Vertex sequences			
No.	Sequence ($5' \rightarrow 3'$)	Tm	GC%
0	AGGCGAGTATGGGGTATATC	60.73	50
1	CCTGTCAACATTGACGCTCA	59.24	50
2	TTATGATTCCACTGGCGCTC	59.00	50
3	ATCGTACTCATGGTCCCTAC	56.81	50
4	CGCTCCATCCTTGATCGTTT	58.13	50
5	CTTCGCTGCTGATAACCTCA	59.44	50
6	GAGTTAGATGTCACGTCACG	56.97	50
Weight sequences			
Edge cost	Sequence ($5' \rightarrow 3'$)	Tm	GC%
3	ATGATAGATATGTAGATTCC	47.89	30
5	GGATGTGATATCGTTCTTGT	54.62	40
7	GGATTAGCAGTGCCTCAGTT	58.37	50
9	TGGCCACGAAGCCTTCCGTT	64.51	60
11	GAGCTGGCTCCTCATCGCGC	68.88	70

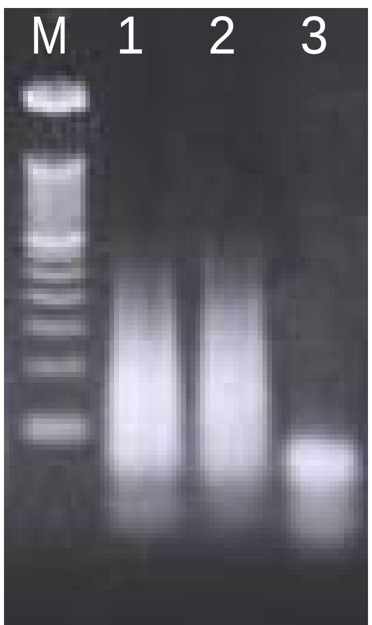
Experimental Implementation for TSP Conditions

- Strating and ending with city 0
  - PCR using primers complementary to city 0
- Visiting every city
  - A series of affinity chromatography
  - Each affinity column contains ssDNA complementary to each city.
- Cheapest path
  - DTG-PCR

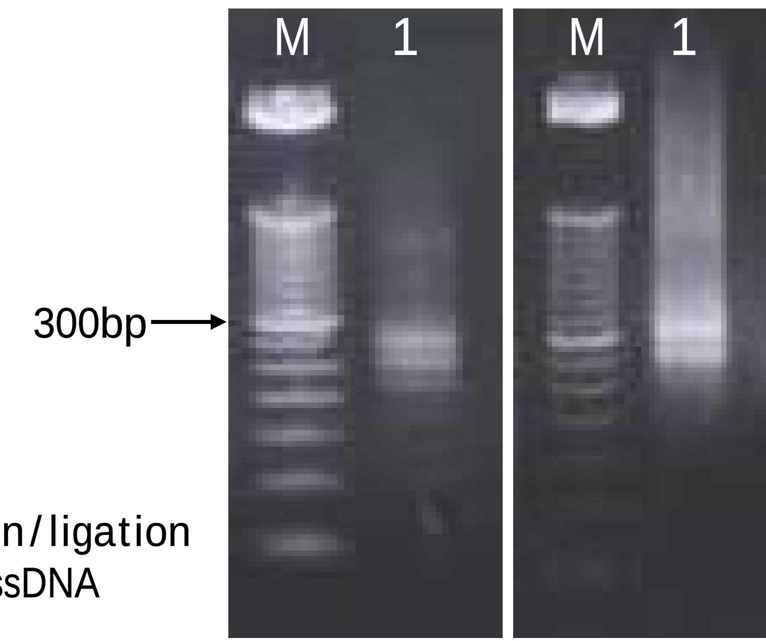
Experimental Results

- Random path generation
  (by hybridization and ligation)

- Selective amplification of paths starting and ending with city 0
  (by PCR using primers complementary to city 0)

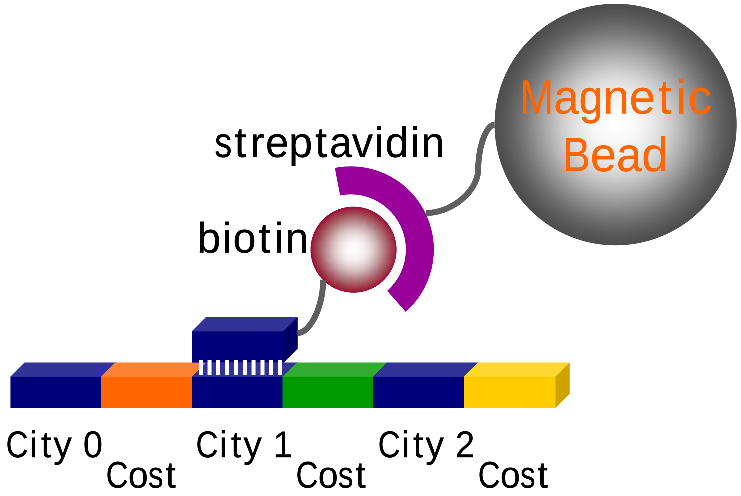


M: 50 bp ladder
lane 1,2:
after hybridization/ligation
lane 3: mixture of ssDNA



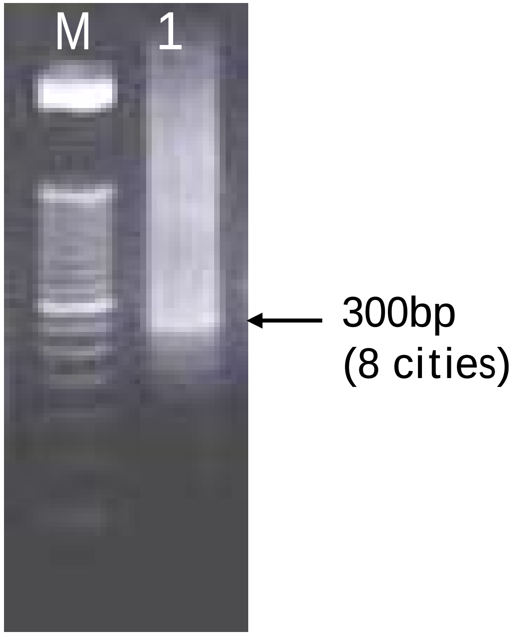
- Separation of paths containing every city

  (by a series of affinity chromatography)

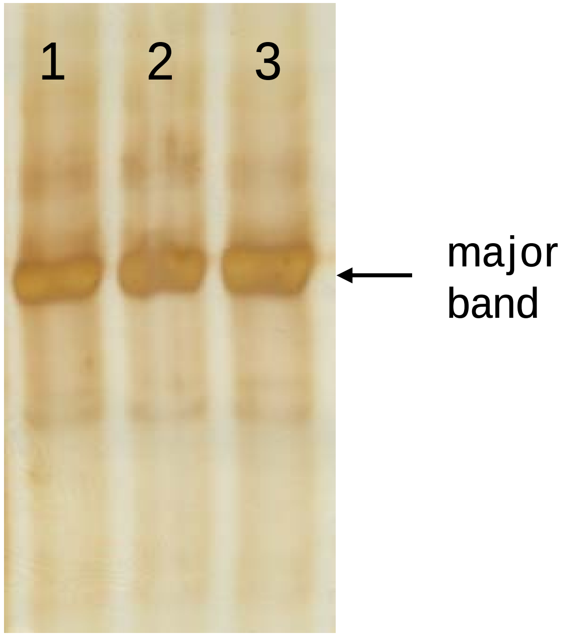


- More amplification of paths with lower costs

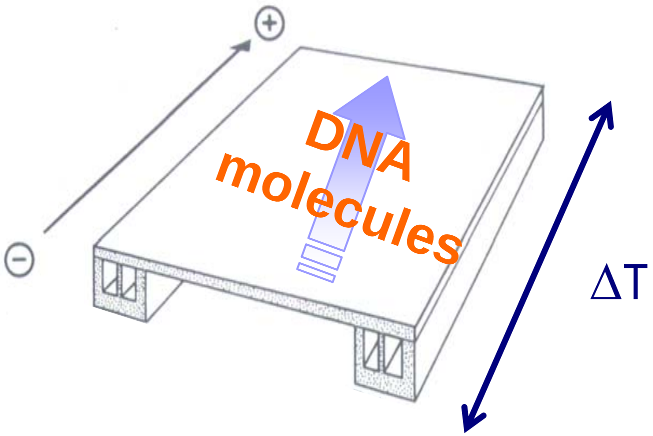
  (by DTG-PCR)



- Separation of the path with lowest cost
  (by TGGE)



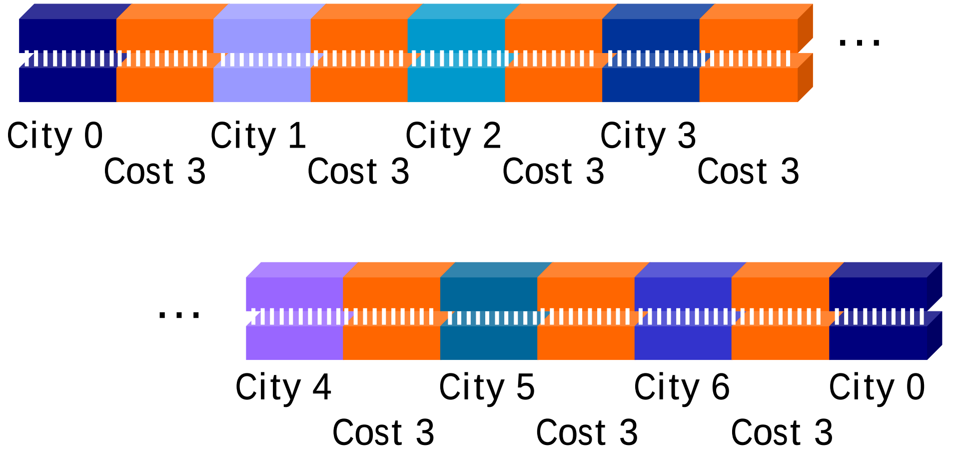
- TGGE
  (Temperature Gradient – Gel Electrophoresis)



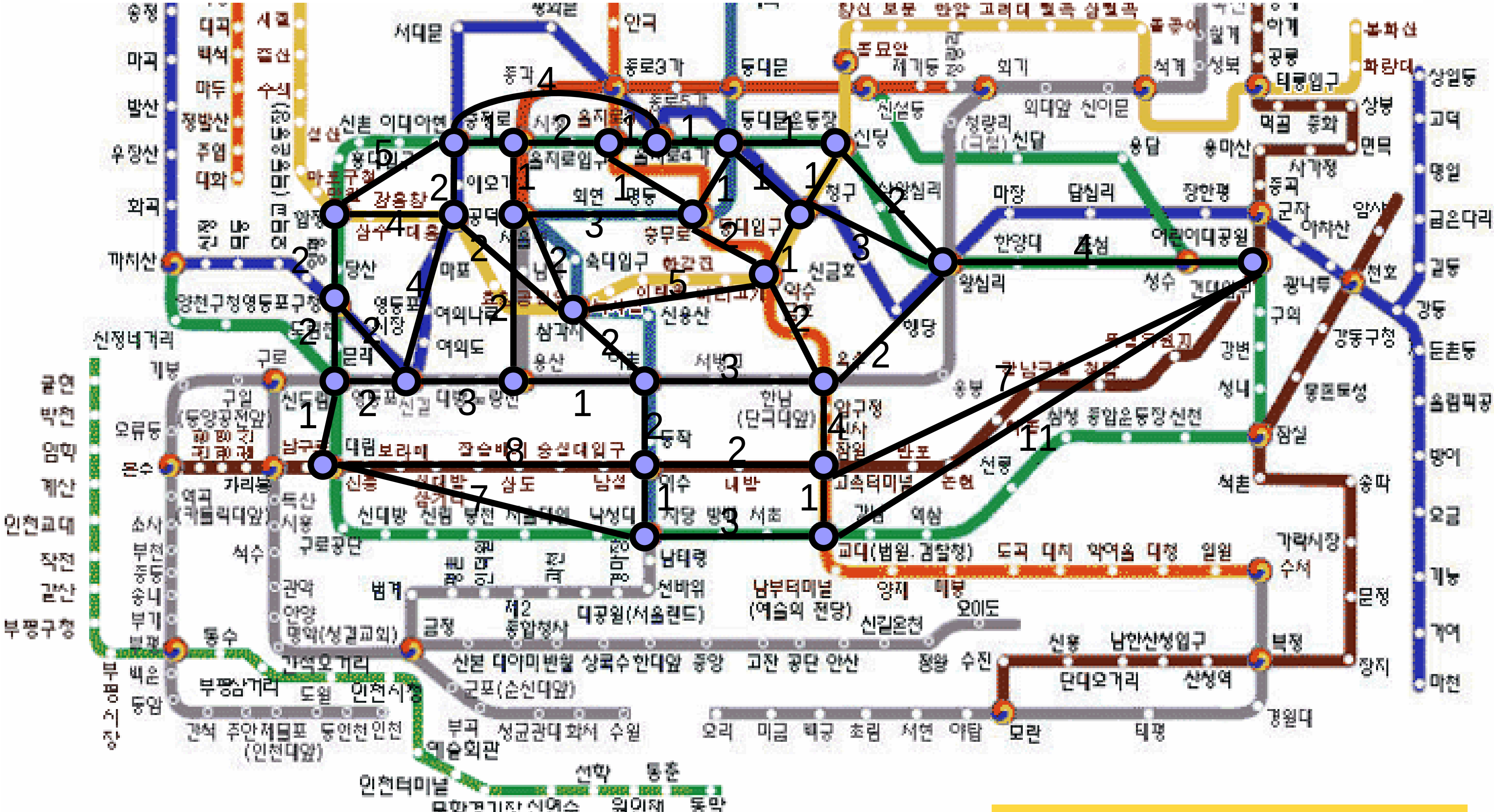
Readout

(by cloning and sequencing)

······TTCTGCGTTGTTTCGGGGTACAGTGGCTCCTCCGTT
CCGCCTGCACTGTGGAGAGGTGAGCAGTGGCTCCTCCGTT
CCGCGTGGATTCACAAGGCCATCGCAGTGGCTCCTCCGTT
CCGCATACGGCGTGGTTTTTCGGGCAGTGGCTCCTCCGTT
CCGCAAACGGTCGTAAGTGATGAACAGTGGCTCCTCCGTT
CCGCGCACAGTCCACCTGTAGACACAGTGGCTCCTCCGTT
CCGCTATGTCCAGCTGTCGCAAAGCAGTGGCTCCTCCGTT
CCGCTTCTGCGTTGTTTCGGGGTA······

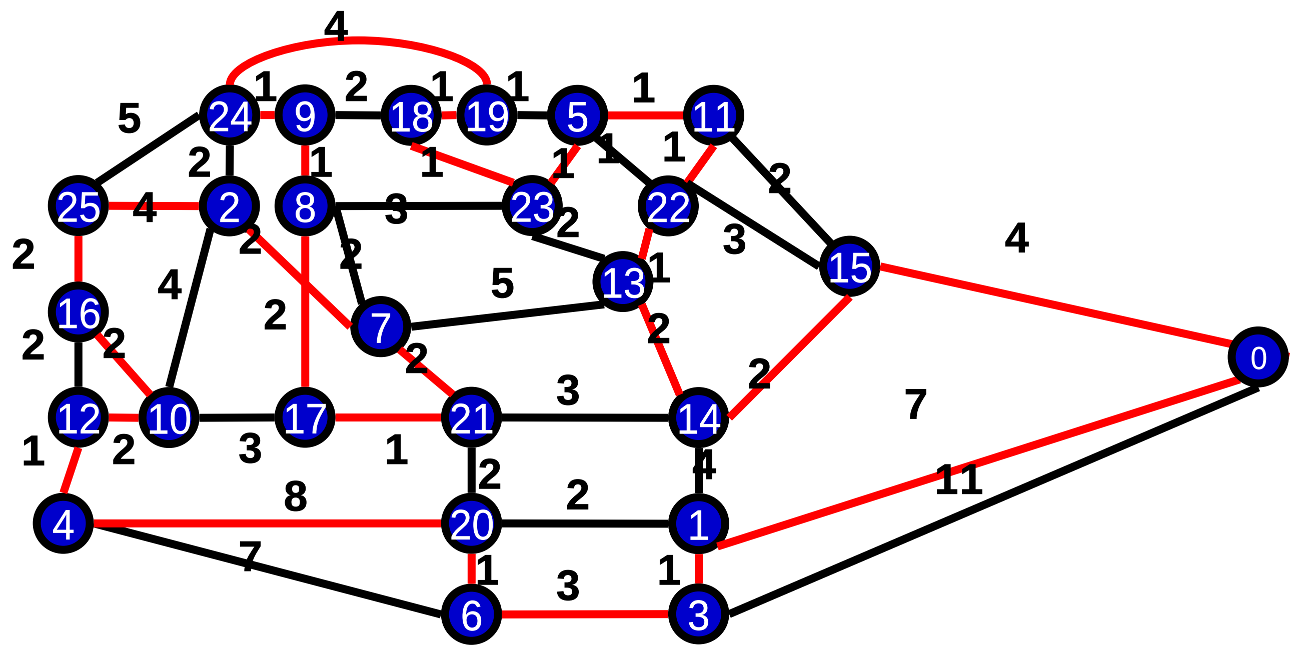


Toward Larger Problems



Subway routes in Seoul

Target Problem: 26-City TSP



- Graph with 26 vertexes (cities) and 92 edges (roads)

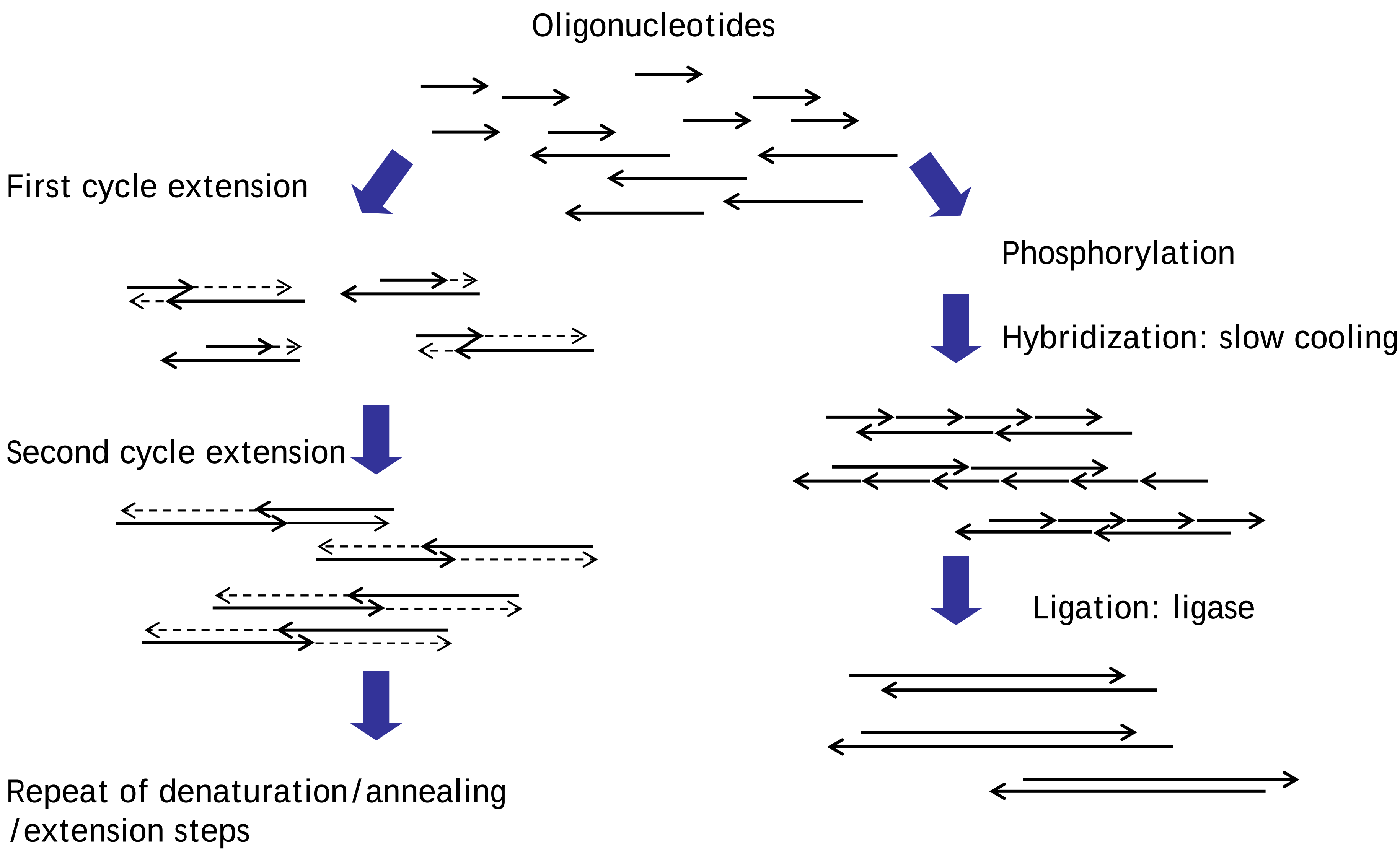
  Vertex: station connected with more than two stations

  Weight: number of stations between vertex stations

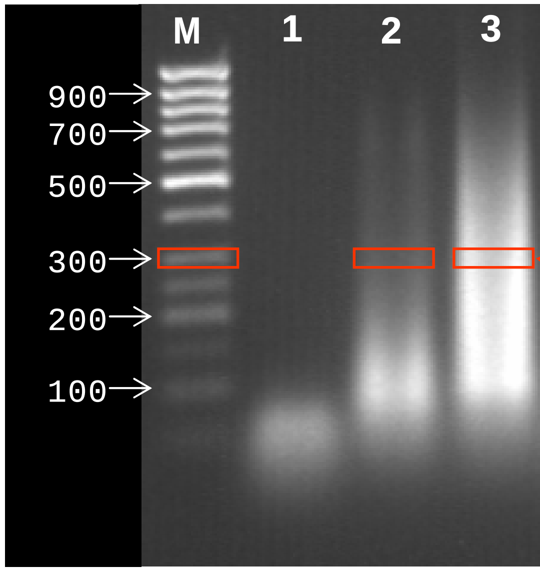
Initial Pool Generation with POA

- Initial pool
  - A combinatorial library that contains numerical or indicative information
- Initial pool generation
  - Prerequisite step of most molecular algorithms for mathematical problems
- Initial pool generation methods
  - Hybridization and ligation method
  - Parallel overlap assembly method

Initial Pool Generation for TSP
: POA vs. Hybiridzation/Ligation

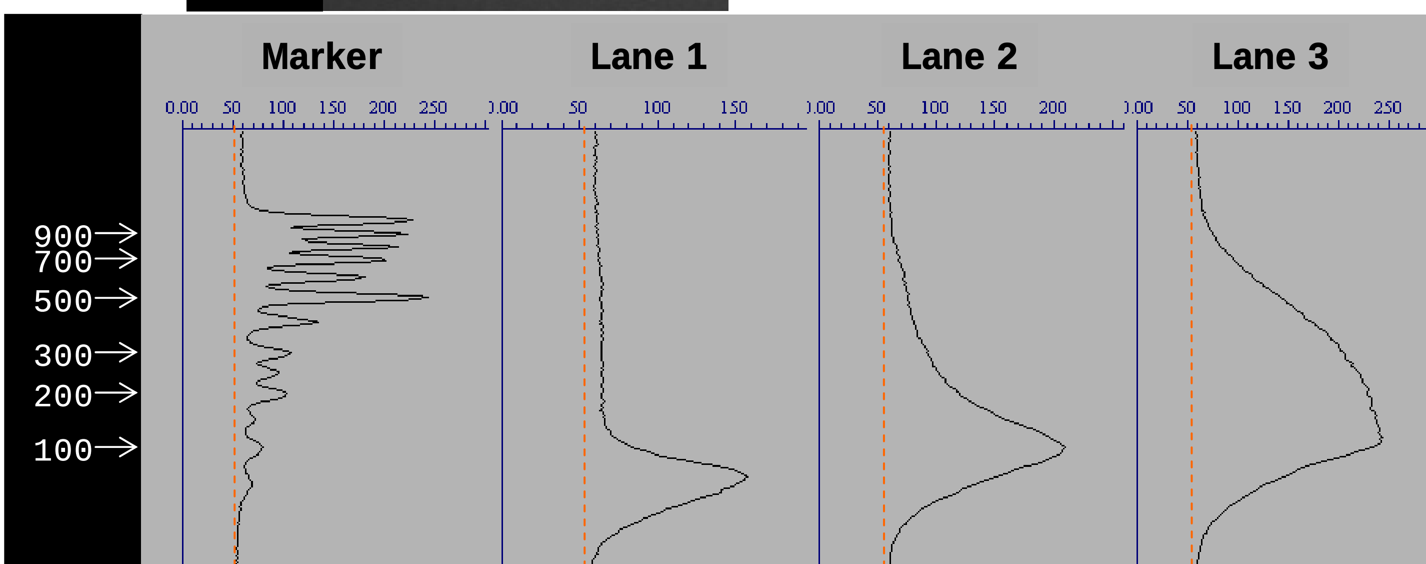


Experimental Results



Lane M: 50 bp ladder
Lane 1: mixture of ssDNA
Lane 2: hybridization/ligation
Lane 3: POA

Quantification by image analysis



Comparison of Two Methods

Parallel overlap assembly	Viewpoint	Hybridization/ligation
Population size is almost preserved throughout the process.	Scalability	Population size decreases as hybridization proceeds.
No need of 5'-phosphorylation	Economy	5'-phosphorylation is necessary
Less time & reagents consuming		More time & reagents consuming
High error rate by non-specific priming	Fidelity	Low error rate (higher hybridization specificity)

Summary

- Development of a molecular algorithm for TSP based on the melting temperature difference
- Development of DTG-PCR as an efficient operator for the graph problem with weighted edge
- Success in solving 7-city TSP
- 26-city traveling salesman problem
- POA as an efficient operator for large size TSP

Acknowledgments

- Prof. Byoung-Tak Zhang

  (School of Computer Science and Engineering, SNU)
- Ji Youn Lee

  (School of Chemical Engineering, SNU)
- Soo-Yong Shin (SNU)

  (School of Computer Science and Engineering, SNU)

- The Ministry of Commerce, Industry, and Energy


'01 9 18